

EGRA Baseline 2013

DQA Report on EGRA baseline student assessments and teacher surveys

Directions: Use the following worksheet to complete an assessment of data for indicators against the 5 data quality standards outlined in the ADS. A comprehensive discussion of each criterion is included in TIPS 12 Data Quality Standards.

Data Quality Assessment Worksheet							
USAID/Uganda Mission							
Development Objective (DO): Increased Literacy and Health Seeking Behaviour							
Intermediate Result (IR): Improved Early Grade Reading and Transition to English							
Indicators: 1a:Proportion of students who, by the end of two grades of primary schooling, demonstrate that they can read and understand the meaning of grade level text in local languages 1.1: Proportion of learners reading at agreed upon benchmark of words per minute (P2,P4) in English and local language, compared to control;1:2: Proportion of learners comprehending at 80% or higher (P2, P4) English and Local Language, compared to control;							
Is this a Standard or Custom Indicator? (Mark "X") Standard X CustomX							
If standard, make sure the title matches the title in the Indicator Handbooks.							
Both EGRA-specific (custom) and Standard Foreign Assistance Indicators are included							
Name of Implementing Partner: Implementing Partner: RTI. Data Quality Assessor: NORC							
Assessment Team Members: NORC Performance and Impact Evaluation team: Varuni Dayaratna, Alicia Menendez, Sarah Hughes, Yvonne Cao, with input from Stella Neema, Evelyn Namubiru, StacyAnn Forrester, Betsy Bassan							
Date (s) of Assessment: Dec 10, 2013							
Data Quality Assessment Methodology: To assess the quality of data collected for the EGRA baseline, NORC's US and Uganda-based team carried out the following activities:							
> Reviewed data collection plans and procedures							
> Reviewed questionnaire content, layout and CAPI functioning							
> Reviewed interviewer training agendas, manuals and planned training pedagogy							
> Observed interviewer training, piloting and data collection							
> Attended data collection debriefing							
> Reviewed raw response datasets for errors in student assessments, teacher/head teacher survey and school inventory							
> Communicated with data cleaning team							
> Reviewed cleaned response datasets for errors in student assessments, teacher/head teacher survey and school inventory							
> Performed baseline analysis of response datasets							
> Reviewed IP's baseline report							
Data & Data Source(s): EGRA student instruments, Head teacher questionnaires, classroom observations							
Is the Indicator Reported to USAID/W?							
Rating: Acceptable Acceptable if Corrections are Made Not Acceptable							
Assessment against the 5 data Quality Standards:							



December 2013

C .:.	D.C		Following (O well Common)
Criterion	Definition	Yes or No	Explanation (Overall Summary)
I. Validity	Do the data clearly and adequately represent the intended result? Some issues to consider are: Face Validity: Would an outsider or an expert in the field agree that the indicator is a valid and logical measure for the stated result? Attribution: Does the indicator measure the contribution of the project? Measurement Error. Are there any measurement errors that could affect the data? Both sampling and non-sampling error should be reviewed.	Unclear	Notes: Face Validity: Literacy and reading skill measures included in the EGRA tool have been used to assess Early Grade Reading in a number of other countries prior to implementation in Uganda. The data collection instruments were adapted to Ugandan language and context via adaptation workshops and pretesting. However, as of the date of this DQA, the implementing partner (IP) has not provided comprehensive information on tests of reliability and validity in its reports. We know from ongoing monitoring and performance evaluation activities that instruments have been customized for each grade and for each language Attribution: not applicable at baseline Measurement Error (Sampling): Cluster I, Round I: At the school level, two schools had to be replaced. At the learner level, RTI did not document any replacements that might have been done. Therefore we are unable to calculate learner response rates. Cluster I Round 2 data collection (November 2013:) Initial reports from the field indicate low response rates (i.e. low numbers of students are being found) in the schools in the Central Region compared to baseline. Schools may require repeat visits to interview students who were unavailable during the first visit, and/or supplementing the treatment sample of 56 schools with a random sub-sample of schools from the baseline of 168 schools that were not selected for Round 2. Data collection procedures for EGRA were standardized and included inperson training as well as piloting of the EGRA instruments prior to data collection among the sample schools. Assessors were monitored (observed) by independent data quality assessors as well as members of MoES and NORC. On the occasions when an assessor was observed veering from established protocols, the data collection supervisory team was alerted to correct the errant behavior.
2. Integrity	Do the data collected, analyzed and reported have established mechanisms in place to reduce manipulation or simple errors in transcription? Note: This criterion requires the reviewer to understand what mechanisms are in place to reduce the possibility of manipulation or transcription error.	Yes	Notes: The data collection tool is programmed as a computer-assisted interview. This mode has been shown to minimize data transcription errors, and NORC's review of the raw and cleaned data show that there are few errors from transcription. Data collection is carried out by the implementing partner. However, the independent evaluator has attended interviewer and supervisor training and observed data collection in the field, providing a level of independent oversight to the data collection.



EGRA Baseline 2013

December 2013

3. Precision	Are data sufficiently precise to present a fair picture of performance and enable management decisionmaking at the appropriate levels?	unclear	Notes: NORC estimated the sample required to detect a double-difference measure of impact of magnitude D = 0.20 with a power of 90%. The final sample for Round I was lower than expected as the data collection team could not always assess 30 students per grade; it was even lower for Round 2 since low numbers of students were being found at the schools. It is therefore unclear at this stage whether the sample size will be sufficiently large to detect impacts.
4. Reliability	Do data reflect stable and consistent data collection processes and analysis methods over time? Note: This criterion requires the reviewer to ensure that the indicator definition is operationally precise (i.e. it clearly defines the exact data to be collected) and to verify that the data is, in fact, collected according to that standard definition consistently over time.	Yes	Notes: The EGRA student assessment tool is programmed as an inperson computer-assisted interview (CAPI) using software that enforces skip patterns and reduces interviewer error (compared to Paper and Pencil). As part of its assessor training, the data collection trainers administer an inter-rater reliability test. The results of the inter-rater reliability tests have not been included in the IP's baseline report. However, the IP invited 100 potential assessors to training and hired the 70 best assessors from among those trained. The data collection processes and analysis methods are documented in writing and are being used to ensure the same procedures are followed in a standardized fashion. Observations by data quality assessors, members of MOeS and NORC's team support ensuring consistency in application of data collection protocols.
5. Timeliness	Are data timely enough to influence management decision-making (i.e. in terms of frequency and currency)?	unclear	Notes: The response data from baseline was fully collected by the end of March 2013. The raw and cleaned data were received by NORC analysts in June, 2013, which was somewhat slower than anticipated. Supporting field and data reports are not complete as of the date of the present DQA, which limits assessors' ability to accurately gauge the quality of the data. In fact, a baseline field report was never produced. However, RTI presented the results of the draft baseline analysis to MOES and USAID stakeholders although it is unclear if this data will be used to inform decision-making at this stage.

A Summary of Key Issues and Recommendations:

Limitations/Key Issues:

1. The Implementing Partner has not provided adequate documentation of the data collection and results of inter-rater reliability tests to assess key quality elements of the data.

Actions Needed to address Limitations/Key Recommendations:

1. IP must report response rates, inter-rater reliability results, and provide a field report for each data collection activity.

Approvals:

	Assessment Team members:	Name: Sarah Hughes Position: Survey Specialist, NORC Name: Yvonne Cao Position: Evaluation Analyst, NORC			
	members.	Ivanie : Ivoline Cao Fosition: Evaluation Analyst, Nonc			
For Office Use Only:					
Team Leader Officer Approval (Office Chief)					



December 2013

Name	
M&E Specialist/SI Advisor/Quality Assurance Specialist (clearance): Name	
Date	